

Komputerowa analiza szeregów czasowych

Temat: Analiza danych rzeczywistych przy pomocy modelu ARMA

Raport należy przygotować w **dwuosobowych grupach** i przesłać przez e-Portal w terminie do **31 stycznia 2025, do końca dnia**. W sprawozdaniu oprócz rezultatów powinny zostać umieszczone również odpowiednie wnioski i interpretacja otrzymanych wyników. Proszę zwrócić uwagę na formę pracy: podział na rozdziały, odpowiednie opracowanie i przedstawienie tabel oraz wykresów itd. Oprócz wyników proszę przedstawić niezbędne definicje, wzory, objaśnienia.

Poniżej propozycja podziału pracy na rozdziały wraz z informacją o niezbędnych do wykonania krokach.

1. Wstęp

- określenie celu pracy,
- informacja o analizowanych danych (czego dane dotyczą, jak długa jest próba, skąd dane pochodzą),
- wizualizacja danych.

2. Przygotowanie danych do analizy

- zbadanie jakości danych (detekcja wartości spoza zakładanego przedziału, identyfikacja ewentualnych braków w danych lub błędów w próbkowaniu),
- (*opcjonalnie*) wyodrębnienie z danych obserwacji do zbioru testowego,
- dekompozycja szeregu czasowego
 - wykres ACF oraz PACF dla surowych danych,
 - (*opcjonalnie*) test ADF weryfikujący hipotezę o niestacjonarności dla surowych danych (Augmented Dickey-Fuller Test),
 - identyfikacja trendów deterministycznych: metody omawiane na wykładzie oraz na laboratorium - (zadanie 4, lista 5) / różnicowanie / różnicowanie sezonowe / transformacje stabilizujące wariancję (Boxa-Coxa),
 - wykres ACF oraz PACF dla uzyskanego szeregu,
 - (*opcjonalnie*) test ADF weryfikujący hipotezę o niestacjonarności dla uzyskanego szeregu (Augmented Dickey-Fuller Test).

3. Modelowanie danych przy pomocy ARMA

- dobranie rzędu modelu (kryteria informacyjne),
- estymacja parametrów modelu wybraną metodą.

4. Ocena dopasowania modelu

- przedziały ufności dla PACF/ACF (zadanie 2 z Listy 6 oraz zadanie 3 z Listy 7),
- porównanie linii kwantylowych z trajektorią,
- (*opcjonalnie*) prognoza dla przyszłych obserwacji i porównanie z rzeczywistymi danymi.

5. Weryfikacja założeń dotyczących szumu

- założenie dot. średniej (wykres wartości resztowych / t test),
- założenie dot. wariancji (wykres wartości resztowych / Modified Levene Test / Arch Test),
- założenie dot. niezależności (wykres ACF/PACF dla wartości resztowych / test Ljunga-Boxa),
- założenie dot. normalności rozkładu (dystrybuenta / gęstość / wykres kwantylowy / testy na normalność).

6. Zakończenie

- podsumowanie pracy,
- wnioski.

Dodatkowe uwagi:

- Proszę zwrócić uwagę na odpowiedni styl i właściwą formę raportu: podział na rozdziały/sekcje, a także na odpowiednie opracowanie i przedstawienie tabeli oraz wykresów. Do każdej tabeli oraz każdego wykresu powinno znaleźć się odwołanie w tekście. Należy je również odpowiednio sformatować (zadbać o to, by wykresy i tabele były czytelne oraz zawierały podpisy).
- Raport powinien zawierać opis danych (informację czego dane dotyczą oraz jak długa jest próba). W pracy należy wskazać skąd pochodzą dane (jeśli jest to możliwe, dobrze jest umieścić w pracy link do strony, z której można pobrać dane).
- Jeżeli korzystają Państwo z jakiś statystyk, wzorów czy twierdzeń to proszę je zamieścić w pracy (oraz odpowiednio zacytować). Np. w przypadku testów statystycznych należy opisać na czym polega test, wskazać hipotezę zerową i hipotezę alternatywną oraz zamieścić wzór statystyki testowej.
- W przypadku badania normalności rozkładu poza testami wizualnymi należy wykonać również testy statystyczne.
- W pracy oprócz rezultatów powinny zostać umieszczone również odpowiednie wnioski i interpretacja otrzymanych wyników – dotyczy to każdej części raportu dot. modelowania i analizy danych (rozdz. 2-5).